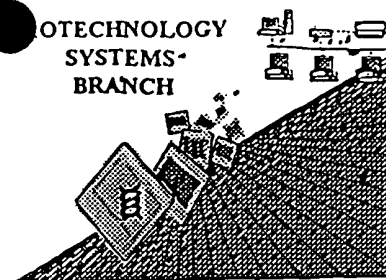Re-run

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/936852_

Source: _PCT09_

Date Processed by STIC: _10/03/01_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

<u>Checker Version 3.0 can be down loaded from the USPTO website at the following address:</u>
http://www.uspto.gov/web/offices/pac/checker

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/936 852

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO

1____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3____ Misaligned Amino
Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5____ Variable Length.

Sequence(s)____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6____ PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7____ Skipped Sequences
(OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8____ Skipped Sequences
(NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence
<210> sequence id number
<400> sequence id number
000

9__✓__ Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10____ Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11____ Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12____ PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

Attention: **Delete End of file NonASCII text

PCT09

RAW SEQUENCE LISTING   DATE: 10/03/2001
PATENT APPLICATION: US/09/936,852   TIME: 10:30:33

Input Set : A:\GKS101-0.txt
Output Set: N:\CRF3\10032001\I936852.raw

```
    3 <110> APPLICANT: Biosyn Arzneimittel GmbH
    5 <120> TITLE OF INVENTION: Nucleic acid molecule comprising a nucleic acid sequence
which
    6       codes for a haemocyanin
    8 <130> FILE REFERENCE: PCT1153-01966
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/936,852
C--> 11 <141> CURRENT FILING DATE: 2001-09-18
   13 <160> NUMBER OF SEQ ID NOS: 108
   15 <170> SOFTWARE: PatentIn Ver. 2.1
```

Does Not Comply
Corrected Diskette Needed

ERRORED SEQUENCES

```
  411 <210> SEQ ID NO: 15
  412 <211> LENGTH: 1546
  413 <212> TYPE: DNA
  414 <213> ORGANISM: Haliotis tuberculata
  416 <400> SEQUENCE: 15
  417 agcccacaga ggaccagttg aagaaacaga agtcactcgc caacatactg acggcaatgc 60
  418 acactttcat cgtaaggaag ttgattcgct gtccctggat gaagcaaaca acttgaagaa 120
  419 tgcccttttac aagctacaga acgaccacag tctaacggga tacgaagcaa tctctggtta 180
  420 ccatggatac cccaatctgt gtccggaaga aggcgatgac aaaataccccc tgctgcgtcc 240
  421 ccggatgggc atctttcctt actggcacag actcttgacc attcaactgg aaagagctct 300
  422 tgagcacaat ggtgcactgc ttggtgttcc ttactgggac tggaacaagg acctgtcgtc 360
  423 actgccggcg ttcttctccg actccagcaa caacaatccc tacttcaagt accacatcgc 420
  424 cggtgttggt cacgacaccg tcagagagcc aactagtctt atatataacc agccccaaat 480
  425 ccatggttat gattatctct attacctagc attgaccacg cttgaagaaa acaattactg 540
  426 ggactttgag gttcagtatg agatcctcca caacgccgtc cactcctggc ttggaggatc 600
  427 ccagaagtat tccatgtcta ccctggagta ttcggccttt gaccctgtct ttatgatcct 660
  428 tcactcgggt ctagacagac tttggatcat ctggcaagaa cttcagaaga tcaggagaaa 720
  429 gccctacaac ttcgctaaat gtgcttatca tatgatggaa gagccactgg cgcccttcag 780
  430 ctatccatct atcaaccagg acgagttcac ccgtgccaac tccaagcctt ctacagtttt 840
  431 tgacagccat aagttcggct accattacga taacctgaat gttagaggtc acagcatcca 900
  432 agaactcaac acaatcatca atgacttgag aaacacagac agaatctacg caggatttgt 960
  433 tttgtcaggc atcggtacgt ctgctagtgt caagatctat ctccgaacag atgacaatga 1020
  434 cgaagaagtt ggaactttca ctgtcctggg aggagagagg gaaatgccat gggcctacga 1080
  435 gcgagttttc aagtatgaca tcacagaggt tgcagataga cttaaaatta agttatgggg 1140
  436 acaccctta acttccggaa ctggagatca catccttacg aatggaatcg gtggtaaaca 1200
  437 agagcctacc caaatccttt catcatctac agacctgcca atcatgacta cgatgttctt 1260
E--> 438 gttatcccag tanggaagaa accttcacat ccctcccaaa gttgtcgtca agaaaggcac 1320
  439 ccgcatcgag ttccacccag tcgatgattc agttacgaga ccagttgttg atcttggaag 1380
  440 ctacactgca ctcttcaact gtgtggtacc accgttcaca taccacggat tcgaactgaa 1440
  441 ccacgtctat tctgtcaagc ctggtgacta ctatgttact ggacccacga gagcctttg 1500
  442 ccagaatgca gatgtcagga ttcatatcca tgttgaggat gagtaa    1546
 1854 <210> SEQ ID NO: 39
 1855 <211> LENGTH: 515
 1856 <212> TYPE: PRT
 1857 <213> ORGANISM: Haliotis tuberculata
```

Must enumerate unknowns in fields 221, 222, and 223.

The type of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.
as well as other

```
1859 <400> SEQUENCE: 39
1860 Gly Ala His Arg Gly Pro Val Glu Glu Thr Glu Val Thr Arg Gln His
1861  1               5                  10                 15
1863 Thr Asp Gly Asn Ala His Phe His Arg Lys Glu Val Asp Ser Leu Ser
1864             20                  25                 30
1866 Leu Asp Glu Ala Asn Asn Leu Lys Asn Ala Leu Tyr Lys Leu Gln Asn
1867         35                  40                 45
1869 Asp His Ser Leu Thr Gly Tyr Glu Ala Ile Ser Gly Tyr His Gly Tyr
1870     50                  55                 60
1872 Pro Asn Leu Cys Pro Glu Glu Gly Asp Asp Lys Ile Pro Leu Leu Arg
1873  65  •              70                 75                 80
1875 Pro Arg Met Gly Ile Phe Pro Tyr Trp His Arg Leu Leu Thr Ile Gln
1876             85                  90                 95
1878 Leu Glu Arg Ala Leu Glu His Asn Gly Ala Leu Leu Gly Val Pro Tyr
1879         100                 105                110
1882 Trp Asp Trp Asn Lys Asp Leu Ser Ser Leu Pro Ala Phe Phe Ser Asp
1883         115                 120                125
1885 Ser Ser Asn Asn Asn Pro Tyr Phe Lys Tyr His Ile Ala Gly Val Gly
1886     130                 135                140
1888 His Asp Thr Val Arg Glu Pro Thr Ser Leu Ile Tyr Asn Gln Pro Gln
1889 145                 150                155                160
1891 Ile His Gly Tyr Asp Tyr Leu Tyr Tyr Leu Ala Leu Thr Thr Leu Glu
1892             165                 170                175
1894 Glu Asn Asn Tyr Trp Asp Phe Glu Val Gln Tyr Glu Ile Leu His Asn
1895             180                 185                190
1897 Ala Val His Ser Trp Leu Gly Gly Ser Gln Lys Tyr Ser Met Ser Thr
1898         195                 200                205
1900 Leu Glu Tyr Ser Ala Phe Asp Pro Val Phe Met Ile Leu His Ser Gly
1901     210                 215                220
1903 Leu Asp Arg Leu Trp Ile Ile Trp Gln Glu Leu Gln Lys Ile Arg Arg
1904 225                 230                235                240
1906 Lys Pro Tyr Asn Phe Ala Lys Cys Ala Tyr His Met Met Glu Glu Pro
1907             245                 250                255
1909 Leu Ala Pro Phe Ser Tyr Pro Ser Ile Asn Gln Asp Glu Phe Thr Arg
1910             260                 265                270
1912 Ala Asn Ser Lys Pro Ser Thr Val Phe Asp Ser His Lys Phe Gly Tyr
1913         275                 280                285
1915 His Tyr Asp Asn Leu Asn Val Arg Gly His Ser Ile Gln Glu Leu Asn
1916     290                 295                300
1918 Thr Ile Ile Asn Asp Leu Arg Asn Thr Asp Arg Ile Tyr Ala Gly Phe
1919 305                 310                315                320
1921 Val Leu Ser Gly Ile Gly Thr Ser Ala Ser Val Lys Ile Tyr Leu Arg
1922             325                 330                335
1924 Thr Asp Asp Asn Asp Glu Glu Val Gly Thr Phe Thr Val Leu Gly Gly
1925             340                 345                350
1927 Glu Arg Glu Met Pro Trp Ala Tyr Glu Arg Val Phe Lys Tyr Asp Ile
1928         355                 360                365
1930 Thr Glu Val Ala Asp Arg Leu Lys Ile Lys Leu Trp Gly His Pro Leu
1931         370                 375                380
```

RAW SEQUENCE LISTING   DATE: 10/03/2001
PATENT APPLICATION: US/09/936,852   TIME: 10:30:33

Input Set : A:\GKS101-0.txt
Output Set: N:\CRF3\10032001\I936852.raw

```
       1933 Thr Ser Gly Thr Gly Asp His Ile Leu Thr Asn Gly Ile Gly Gly Lys
       1934 385             390             395             400
       1936 Gln Glu Pro Thr Gln Ile Leu Ser Ser Ser Thr Asp Leu Pro Ile Met
       1937         405             410             415
E-->   1939 Thr Thr Met Phe Leu Leu Ser Gln Xaa Gly Arg Asn Leu His Ile Pro
       1940         420             425             430
       1942 Pro Lys Val Val Val Lys Lys Gly Thr Arg Ile Glu Phe His Pro Val
       1943         435             440             445
       1945 Asp Asp Ser Val Thr Arg Pro Val Val Asp Leu Gly Ser Tyr Thr Ala
       1946     450             455             460
       1948 Leu Phe Asn Cys Val Val Pro Pro Phe Thr Tyr His Gly Phe Glu Leu
       1949 465             470             475             480
       1951 Asn His Val Tyr Ser Val Lys Pro Gly Asp Tyr Tyr Val Thr Gly Pro
       1952             485             490             495
       1954 Thr Arg Asp Leu Cys Gln Asn Ala Asp Val Arg Ile His Ile His Val
       1955         500             505             510
       1957 Glu Asp Glu
       1958         515
       2851 <210> SEQ ID NO: 62
       2852 <211> LENGTH: 1185
       2853 <212> TYPE: DNA
       2854 <213> ORGANISM: Haliotis tuberculata
       2856 <400> SEQUENCE: 62
       2857 atcatattgc tggcagtgga gtcaggaaag acgtgacgtc tcttaccgca tctgagatag 60
       2858 agaacctgag gcatgctctg caaagcgtga tggatgatga tggacccaat ggattccagg 120
E-->   2859 caattgctgc ttatcacgga agtcctccca tgtgtcacat gcntgatggt agagacgttg 180
       2860 catgttgtac tcatggaatg gcatcttttcc ctcactggca cagactgttt gtgaaacaga 240
       2861 tggaggatgc actggctgcg catggagctc acattggcat accatactgg gattggacaa 300
       2862 gtgcgtttag tcatctgcct gccctagtga ctgaccacga gcacaatccc ttccaccacg 360
       2863 gacatattgc tcatcggaat gtggatacat ctcgatctcc gagagacatg ctgttcaatg 420
       2864 accccgaaca cgggtcagaa tcattcttct atagacaggt tctcttggct ctagaacaga 480
       2865 cagacttctg ccaatttgaa gttcagtttg aaataacaca caatgcaatc cactcttgga 540
       2866 ctggaggaca tactccatat ggaatgtcat cactggaata tacagcatat gatccactct 600
       2867 tttatctcca ccattccaac actgatcgta tctgggccat ctggcaggca ctccagaaat 660
       2868 acagaggttt tcaatacaac gcagctcatt gcgatatcca ggttctgaaa caacctctta 720
       2869 aaccattcag cgagtccagg aatccaaacc cagtcaccag agccaattct agggcagtcg 780
       2870 attcatttga ttatgagaga ctcaattatc aaatatgacac acttaccttc cacggacatt 840
       2871 ctatctcaga acttgatgcc atgcttcaag agagaagaa ggaagagaga acatttgcag 900
       2872 ccttcctgtt gcacggattt ggcgccagtg ctgatgtttc gtttgatgtc tgcacacctg 960
       2873 atggtcattg tgcctttgct ggaaccttcg cggtacttgg tggggagctt gagatgccct 1020
       2874 ggtcctttga aagattgttc cgttacgata tcacaaaggt tctcaagcag atgaatcttc 1080
       2875 actatgattc tgagttccac tttgagttga agattgttgg cacagatgga acagaactgc 1140
       2876 catcggatcg tatcaagagc cctaccattg aacaccatgg aggag 1185
       4136 <210> SEQ ID NO: 79
       4137 <211> LENGTH: 395
       4138 <212> TYPE: PRT
       4139 <213> ORGANISM: Megathura crenulata
       4141 <400> SEQUENCE: 79
       4142 Asp His Ile Ala Gly Ser Gly Val Arg Lys Asp Val Thr Ser Leu Thr
```

Must enumerate unknowns

Must enumerate unknowns

```
      4143   1                5                       10                      15
      4145 Ala Ser Glu Ile Glu Asn Leu Arg His Ala Leu Gln Ser Val Met Asp
      4146               20                      25                      30
      4148 Asp Asp Gly Pro Asn Gly Phe Gln Ala Ile Ala Ala Tyr His Gly Ser
      4149               35                      40                      45
E--> 4151 Pro Pro Met Cys His Met Xaa Asp Gly Arg Asp Val Ala Cys Cys Thr
      4152        50                      55                      60
      4154 His Gly Met Ala Ser Phe Pro His Trp His Arg Leu Phe Val Lys Gln
      4155  65                      70                      75                      80
      4157 Met Glu Asp Ala Leu Ala Ala His Gly Ala His Ile Gly Ile Pro Tyr
      4158               85                      90                      95
      4160 Trp Asp Trp Thr Ser Ala Phe Ser His Leu Pro Ala Leu Val Thr Asp
      4161              100                     105                     110
      4163 His Glu His Asn Pro Phe His His Gly His Ile Ala His Arg Asn Val
      4164              115                     120                     125
      4166 Asp Thr Ser Arg Ser Pro Arg Asp Met Leu Phe Asn Asp Pro Glu His
      4167        130                     135                     140
      4169 Gly Ser Glu Ser Phe Phe Tyr Arg Gln Val Leu Leu Ala Leu Glu Gln
      4170 145                     150                     155                     160
      4172 Thr Asp Phe Cys Gln Phe Glu Val Gln Phe Glu Ile Thr His Asn Ala
      4173                   165                     170                     175
      4175 Ile His Ser Trp Thr Gly Gly His Thr Pro Tyr Gly Met Ser Ser Leu
      4176              180                     185                     190
      4178 Glu Tyr Thr Ala Tyr Asp Pro Leu Phe Tyr Leu His His Ser Asn Thr
      4179              195                     200                     205
      4181 Asp Arg Ile Trp Ala Ile Trp Gln Ala Leu Gln Lys Tyr Arg Gly Phe
      4182        210                     215                     220
      4184 Gln Tyr Asn Ala Ala His Cys Asp Ile Gln Val Leu Lys Gln Pro Leu
      4185 225                     230                     235                     240
      4187 Lys Pro Phe Ser Glu Ser Arg Asn Pro Asn Pro Val Thr Arg Ala Asn
      4188              245                     250                     255
      4190 Ser Arg Ala Val Asp Ser Phe Asp Tyr Glu Arg Leu Asn Tyr Gln Tyr
      4191              260                     265                     270
      4193 Asp Thr Leu Thr Phe His Gly His Ser Ile Ser Glu Leu Asp Ala Met
      4194        275                     280                     285
      4196 Leu Gln Glu Arg Lys Lys Glu Glu Arg Thr Phe Ala Ala Phe Leu Leu
      4197        290                     295                     300
      4199 His Gly Phe Gly Ala Ser Ala Asp Val Ser Phe Asp Val Cys Thr Pro
      4200 305                     310                     315                     320
      4202 Asp Gly His Cys Ala Phe Ala Gly Thr Phe Ala Val Leu Gly Gly Glu
      4203                   325                     330                     335
      4205 Leu Glu Met Pro Trp Ser Phe Glu Arg Leu Phe Arg Tyr Asp Ile Thr
      4206        340                      345                     350
      4208 Lys Val Leu Lys Gln Met Asn Leu His Tyr Asp Ser Glu Phe His Phe
      4209        355                     360                     365
      4211 Glu Leu Lys Ile Val Gly Thr Asp Gly Thr Glu Leu Pro Ser Asp Arg
      4212        370                     375                     380
      4214 Ile Lys Ser Pro Thr Ile Glu His His Gly Gly
      4215 385                     390                     395
```

```
 4636 <210> SEQ ID NO: 95
 4637 <211> LENGTH: 1548
 4638 <212> TYPE: DNA
 4639 <213> ORGANISM: Haliotis tuberculata
 4641 <400> SEQUENCE: 95
 4642 ggagcccaca gaggaccagt tgaagaaaca gaagtcactc gccaacatac tgacggcaat 60
 4643 gcacactttc atcgtaagga agttgattcg ctgtccctgg atgaagcaaa caacttgaag 120
 4644 aatgcccttt acaagctaca gaacgaccac agtctaacgg gatacgaagc aatctctggt 180
 4645 taccatggat accccaatct gtgtccggaa gaaggcgatg acaaaatacc cctgctgcgt 240
 4646 ccccggatgg gcatctttcc ttactggcac agactcttga ccattcaact ggaaagagct 300
 4647 cttgagcaca atggtgcact gcttggtgtt ccttactggg actggaacaa ggacctgtcg 360
 4648 tcactgccgg cgttcttctc cgactccagc aacaacaatc cctacttcaa gtaccacatc 420
 4649 gccggtgttg gtcacgacac cgtcagagag ccaactagtc ttatatataa ccagccccaa 480
 4650 atccatggtt atgattatct ctattaccta gcattgacca cgcttgaaga aaacaattac 540
 4651 tgggactttg aggttcagta tgagatcctc cacaacgccg tccactcctg gcttggagga 600
 4652 tcccagaagt attccatgtc taccctggag tattcggcct ttgaccctgt ctttatgatc 660
 4653 cttcactcgg gtctagacag actttggatc atctggcaag aacttcagaa gatcaggaga 720
 4654 aagccctaca acttcgctaa atgtgcttat catatgatgg aagagccact ggcgcccttc 780
 4655 agctatccat ctatcaacca ggacgagttc acccgtgcca actccaagcc ttctacagtt 840
 4656 tttgacagcc ataagttcgg ctaccattac gataacctga atgttagagg tcacagcatc 900
 4657 caagaactca acacaatcat caatgacttg agaaacacag acagaatcta cgcaggattt 960
 4658 gttttgtcag gcatcggtac gtctgctagt gtcaagatct atctccgaac agatgacaat 1020
 4659 gacgaagaag ttggaacttt cactgtcctg ggaggagaga gggaaatgcc atgggcctac 1080
 4660 gagcgagttt tcaagtatga catcacagag gttgcagata gacttaaaat taagttatgg 1140
 4661 ggacacccct taacttccgg aactggagat cacatcctta cgaatggaat cggtggtaaa 1200
 4662 caagagccta cccaaatcct ttcatcatct acagacctgc caatcatgac tacgatgttc 1260
E--> 4663 ttgttatccc agtangggaag aaaccttcac atccctccca aagttgtcgt caagaaaggc 1320
 4664 acccgcatcg agttccaccc agtcgatgat tcagttacga gaccagttgt tgatcttgga 1380
 4665 agctacactg cactcttcaa ctgtgtggta ccaccgttca cataccacgg attcgaactg 1440
 4666 aaccacgtct attctgtcaa gcctggtgac tactatgtta ctggacccac gagagacctt 1500
 4667 tgccagaatg cagatgtcag gattcatatc catgttgagg atgagtaa        1548
 4960 <210> SEQ ID NO: 107
 4961 <211> LENGTH: 1185
 4962 <212> TYPE: DNA
 4963 <213> ORGANISM: Megathura crenulata
 4965 <400> SEQUENCE: 107
 4966 gatcatattg ctggcagtgg agtcaggaaa gacgtgacgt ctcttaccgc atctgagata 60
 4967 gagaacctga ggcatgctct gcaaagcgtg atggatgatg atggacccaa tggattccag 120
E--> 4968 gcaattgctg cttatcacgg aagtcctccc atgtgtcaca tgcntgatgg tagagacgtt 180
 4969 gcatgttgta ctcatggaat ggcatctttc cctcactggc acagactgtt tgtgaaacag 240
 4970 atggaggatg cactggctgc gcatggagct cacattggca taccatactg ggattggaca 300
 4971 agtgcgttta gtcatctgcc tgccctagtg actgaccacg agcacaatcc cttccaccac 360
 4972 ggacatattg ctcatcggaa tgtggataca tctcgatctc cgagagacat gctgttcaat 420
 4973 gaccccgaac acgggtcaga atcattcttc tatagacagg ttctcttggc ctagaacag 480
 4974 acagacttct gccaatttga agttcagttt gaaataacac acaatgcaat ccactcttgg 540
 4975 actggaggac atactccata tggaatgtca tcactggaat atacagcata tgatccactc 600
 4976 tttttatctcc accattccaa cactgatcgt atctgggcca tctggcaggc actccagaaa 660
 4977 tacagaggtt ttcaatacaa cgcagctcat tgcgatatcc aggttctgaa acaacctctt 720
 4978 aaaccattca gcgagtccag gaatccaaac ccagtcacca gagccaattc tagggcagtc 780
```

```
4979 gattcatttg attatgagag actcaattat caatatgaca cacttacctt ccacggacat 840
4980 tctatctcag aacttgatgc catgcttcaa gagagaaaga aggaagagag aacatttgca 900
4981 gccttcctgt tgcacggatt tggcgccagt gctgatgttt cgtttgatgt ctgcacacct 960
4982 gatggtcatt gtgcctttgc tggaaccttc gcggtacttg gtggggagct tgagatgccc 1020
4983 tggtcctttg aaagattgtt ccgttacgat atcacaaagg ttctcaagca gatgaatctt 1080
4984 cactatgatt ctgagttcca ctttgagttg aagattgttg gcacagatgg aacagaactg 1140
4985 ccatcggatc gtatcaagag ccctaccatt gaacaccatg gagga              1185
4987 <210> SEQ ID NO: 108
4988 <211> LENGTH: 309
4989 <212> TYPE: DNA
4990 <213> ORGANISM: Megathura crenulata
4992 <400> SEQUENCE: 108
4993 ggtcacgatc acagtgaacg tcacgatgga tttttcagga aggaagtcgg ttccctgtcc 60
4994 ctggatgaag ccaatgacct taaaaatgca ctgtacaagc tgcagaatga tcagggtccc 120
4995 aatggatatg aatcaatagc cggttaccat ggctatccat tcctctgccc tgaacatggt 180
4996 gaagaccagt acgcatgctg tgtccacgga atgcctgtat ttccacattg gcacagactt 240
4997 catacaatcc agtttgagag agctctcaaa gaacatggtt ctcatttggg tctgccatac 300
4998 tgggactgg                                               309
E--> 5003 88
```

Delete NonASCII end of file text (handwritten note)

VERIFICATION SUMMARY   DATE: 10/03/2001
PATENT APPLICATION: **US/09/936,852**   TIME: 10:30:35

Input Set : **A:\GKS101-0.txt**
Output Set: **N:\CRF3\10032001\I936852.raw**

L:10 M:270 C: Current Application Number differs, Replaced Application Number
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:438 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:15
L:1939 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:39
L:2859 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:62
L:4151 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:79
L:4663 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:95
L:4968 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:107
L:5003 M:254 E: No. of Bases conflict, LENGTH:Input:88 Counted:309 SEQ:108